

## Guía docente de la asignatura

<b>Asignatura</b>	TECNOLOGÍA PARA EL BIG DATA		
<b>Materia</b>	BIG DATA		
<b>Módulo</b>	TECNOLOGÍAS INFORMÁTICAS DEL BIG DATA		
<b>Titulación</b>	MÁSTER EN INGENIERÍA INFORMÁTICA		
<b>Plan</b>	510	<b>Código</b>	53203
<b>Periodo de impartición</b>	1 <sup>er</sup> CUATRIMESTRE	<b>Tipo/Carácter</b>	OPTATIVA
<b>Nivel/Ciclo</b>	MÁSTER	<b>Curso</b>	2º
<b>Créditos ECTS</b>	6 ECTS		
<b>Lengua en que se imparte</b>	CASTELLANO		
<b>Profesor/es responsable/s</b>	CÉSAR GONZÁLEZ FERRERAS IVÁN SANTOS TEJIDO		
<b>Datos de contacto (E-mail, teléfono...)</b>	TELÉFONO: 983 185623 - E-MAIL: <a href="mailto:cesargf@infor.uva.es">cesargf@infor.uva.es</a> TELÉFONO: 983 423000 ext. 5512 - E-MAIL: <a href="mailto:ivasan@tel.uva.es">ivasan@tel.uva.es</a>		
<b>Horario de tutorías</b>	Véase <a href="http://www.uva.es/export/sites/uva/2.docencia/2.02.mastersoficiales/2.02.01.ofertaeducativa/2.02.01.01.alfabetica/Ingenieria-Informatica/">http://www.uva.es/export/sites/uva/2.docencia/2.02.mastersoficiales/2.02.01.ofertaeducativa/2.02.01.01.alfabetica/Ingenieria-Informatica/</a>		
<b>Departamento</b>	DEPARTAMENTO DE INFORMÁTICA DEPARTAMENTO DE ELECTRÓNICA		

## 1. Situación / Sentido de la Asignatura

### 1.1 Contextualización

Esta asignatura pretende introducir las diversas tecnologías que se utilizan en las aplicaciones de análisis de grandes repositorios de datos. Se profundizará en las herramientas de análisis distribuido de datos más utilizadas. Por otro lado, se estudiará el proceso de puesta en marcha de la infraestructura necesaria para este tipo de aplicaciones. Se presentarán las diferentes plataformas de computación en la nube existentes, y se analizarán las diversas posibilidades que ofrecen para las aplicaciones big data.

### 1.2 Relación con otras materias

Esta asignatura se relaciona con conceptos presentados en las siguientes asignaturas:

- Técnicas Escalables de Análisis de Datos
- Almacenamiento Escalable
- Big Data: inteligencia de negocios

### 1.3 Prerrequisitos

## 2. Competencias

### 2.1 Generales

### 2.2 Específicas

Código	Descripción
CET1	Capacidad para modelar, diseñar, definir la arquitectura, implantar, gestionar, operar, administrar y mantener aplicaciones, redes, sistemas, servicios y contenidos informáticos.
CET5	Capacidad para analizar las necesidades de información que se plantean en un entorno y llevar a cabo en todas sus etapas el proceso de construcción de un sistema de información.
CET6	Capacidad para diseñar y evaluar sistemas operativos y servidores, y aplicaciones y sistemas basados en computación distribuida.
CET9	Capacidad para aplicar métodos matemáticos, estadísticos y de inteligencia artificial para modelar, diseñar y desarrollar aplicaciones, servicios, sistemas inteligentes y sistemas basados en el conocimiento.
CBD1	Capacidad de aplicar técnicas de análisis escalable de datos en diferentes ámbitos de actuación utilizando las tecnologías adecuadas.
CBD4	Capacidad de implementar sistemas de descubrimiento de conocimiento en grandes bases de datos distribuidas.
CBD5	Capacidad de analizar, diseñar y construir o configurar sistemas de almacenamiento escalable y procesamiento escalable

## 3. Objetivos

Código	Descripción
CBD1.1	Conocer los principales lenguajes de programación utilizados en Big Data y saber programar en detalle con al menos uno de ellos.
CBD4.1	Conocer las herramientas de análisis distribuido de datos que existen junto con sus principales características. Saber manejar y resolver problemas sencillos con alguna de las

	diferentes herramientas de análisis distribuido de datos.
CBD5.1	Conocer los diferentes elementos software que se pueden utilizar para configurar una infraestructura dedicada al Big Data. Saber crear y administrar una infraestructura sencilla de Big Data.
CBD5.2	Conocer las diferentes plataformas de computación en la nube existentes junto con sus características. Desarrollar y configurar proyectos de Big Data utilizando dichas plataformas.

#### 4. Tabla de dedicación del estudiante a la asignatura

ACTIVIDADES PRESENCIALES	HORAS	ACTIVIDADES NO PRESENCIALES	HORAS
Clases teórico-prácticas (T/M)	26	Estudio y trabajo autónomo individual	45
Clases prácticas de aula (A)		Estudio y trabajo autónomo grupal	45
Laboratorios (L)	30		
Prácticas externas, clínicas o de campo			
Seminarios (S)	4		
Tutorías grupales (TG)			
Evaluación (fuera del periodo oficial de exámenes)			
<b>Total presencial</b>	<b>60</b>	<b>Total no presencial</b>	<b>90</b>

#### 5. Bloques temáticos

##### Bloque 1: Lenguajes y Herramientas para el Big Data

Carga de trabajo en créditos ECTS: 

##### a. Contextualización y justificación

En este bloque se trabajan las competencias relacionadas con los lenguajes y las herramientas orientados al big data.

##### b. Objetivos de aprendizaje

Código	Descripción
CBD1.1	Conocer los principales lenguajes de programación utilizados en Big Data y saber programar en detalle con al menos uno de ellos.
CBD4.1	Conocer las herramientas de análisis distribuido de datos que existen junto con sus principales características. Saber manejar y resolver problemas sencillos con alguna de las diferentes herramientas de análisis distribuido de datos.

##### c. Contenidos

**TEMA 1: Lenguajes para el Big Data****TEMA 2: Introducción a la minería de datos**

**TEMA 3: Modelo de Programación MapReduce****TEMA 4: Herramientas para análisis distribuido****d. Métodos docentes**

---

Ver anexo: métodos docentes.

**e. Plan de trabajo**

---

Se desarrollarán dos entregas prácticas empleando las tecnologías desarrolladas en la asignatura.

**f. Evaluación**

---

Ver apartado 7.

**g. Bibliografía básica**

---

Wes McKinney. "Python for Data Analysis". O'Reilly Media, 2012.

Ivan Idris. "Python Data Analysis". Packt Publishing, 2014

Jure Leskovec, Anand Rajaraman, Jeff Ullman "Mining of Massive Datasets" (<http://www.mmds.org/>)

Tom White. "Hadoop: The Definitive Guide". O'Reilly Media / Yahoo Press. 2012

**h. Bibliografía complementaria**

---

**i. Recursos necesarios**

---

- El alumno deberá tener acceso a un ordenador personal.
- Aula virtual de la asignatura.
- Software:
  - o Python
  - o Hadoop

**Bloque 2: Infraestructuras y Plataformas para el Big Data**

---

Carga de trabajo en créditos ECTS:

**a. Contextualización y justificación**

---

En este bloque se trabajan las competencias relacionadas con las infraestructuras y las plataformas para las aplicaciones de big data.

**b. Objetivos de aprendizaje**

---

Código	Descripción
CBD5.1	Conocer los diferentes elementos software que se pueden utilizar para configurar una

	infraestructura dedicada al Big Data. Saber crear y administrar una infraestructura sencilla de Big Data.
CBD5.2	Conocer las diferentes plataformas de computación en la nube existentes junto con sus características. Desarrollar y configurar proyectos de Big Data utilizando dichas plataformas.

### c. Contenidos

**TEMA 5: Infraestructura para el Big Data**

**TEMA 6: Plataformas en la nube**

### d. Métodos docentes

Ver 8. Anexo: Métodos docentes

### e. Plan de trabajo

Ver 9. Anexo: Cronograma de actividades previstas

### f. Evaluación

Ver 7. Sistema de calificaciones – Tabla resumen.

### g. Bibliografía básica

### h. Bibliografía complementaria

### i. Recursos necesarios

El profesor de la asignatura proporcionará los recursos necesarios para realizar este bloque.

## 6. Temporalización (por bloques temáticos)

BLOQUE TEMÁTICO	CARGA ECTS	PERIODO PREVISTO DE DESARROLLO
Bloque 1: Lenguajes y Herramientas para el Big Data	4 ECTS	Semanas 1 a 10
Bloque 2: Infraestructuras y Plataformas para el Big Data	2 ECTS	Semanas 11 a 15

## 7. Sistema de calificaciones – Tabla resumen

INSTRUMENTO/PROCEDIMIENTO	PESO EN LA NOTA FINAL	OBSERVACIONES
Bloque 1: entrega y defensa de la práctica 1	1/3	Semana 5
Bloque 1: entrega y defensa de la práctica 2	1/3	Semana 10
Bloque 3: entrega y defensa de la práctica 3	1/3	Semana 15

**CRITERIOS DE CALIFICACIÓN****Convocatoria extraordinaria:**

- Para la realización de la parte práctica será necesario contactar con el profesor y fijar una planificación para el desarrollo de la misma. En todo caso, será necesario entregar y defender un trabajo práctico.

**8. Anexo: Métodos docentes**

Actividad	Metodología
<b>Clase de teoría</b>	<ul style="list-style-type: none"><li>• Clase magistral participativa</li><li>• Estudio de casos en aula</li><li>• Resolución de problemas</li></ul>
<b>Clase práctica</b>	<ul style="list-style-type: none"><li>• Realización de un trabajo práctico guiado por el profesor.</li><li>• Clase magistral participativa</li><li>• Resolución de casos prácticos con apoyo informático</li><li>• Realización de un proyecto en grupo (2/3 alumnos) guiado por el profesor, siguiendo un enfoque colaborativo.</li></ul>
<b>Seminarios</b>	<ul style="list-style-type: none"><li>• Talleres de aprendizaje</li><li>• Sesiones de debate entre alumnos y profesor sobre su aprendizaje, las técnicas estudiadas y su aplicación práctica a casos reales.</li></ul>

**9. Anexo: Cronograma de actividades previstas**

El cronograma de actividades aparecerá en el Aula Virtual al inicio de la asignatura.